[*Translating and the Computer 26*, November 2004 [London: Aslib, 2004]

Machine Translatability and Post-Editing Effort:

How do they relate?

Sharon O'Brien

School of Applied Language and Intercultural Studies

Dublin City University

Glasnevin

Dublin 9, Ireland.

sharon.obrien@dcu.ie

1. Background

In the early 1990s, translation service providers, in particular those who worked for IT companies, began using translation memory tools in order to meet the ever-growing demand for translation. Translation Memory tools proved to be efficient in reducing the amount of re-work involved in translating documentation and help systems for products that were continually updated by reducing the amount of re-translation and cutting and pasting that were the norm prior to that era.

Now it is 2004 and we still hear reports about the ever-growing demand for translation. According to an IDC survey of the globalisation, internationalisation and localisation market, the annual growth rate is 16.3%. The localisation and translation services part of this market is growing at an annual rate of 14.6% (Van der Meer: 2003). Clearly, this level of growth increases the need for faster translation throughput. In addition, there is a growing demand for "gist" translation, or translation for informational purposes only. For example, members of various EU institutions frequently request gist translations of documents in order to ascertain whether or not the document ought to be translated professionally. Users of web sites in languages other than their mother tongue also avail of gist translation. So, while translation memory tools play an extremely important part in meeting the world-wide need for translation, demand for automatic translation by computers (i.e. machine translation - MT) is also on the increase.

It has been widely acknowledged that when the aim of machine translation is to produce high-quality translation, then post-editing by human translators is necessary. The time and effort required for post-editing can be reduced by implementing a number of strategies. For example, previously translated sentences can be leveraged from a translation memory first, thereby reducing the number of words that have to be machine translated and post-edited. Also, machine translation quality can be improved by adding terms to user-specific MT dictionaries. A third strategy is to use controlled language (CL) to improve the *translatability* of the source text (Mitamura et al. 1998; Nyberg et al. 2003; EAMT/CLAW 2003). This paper focuses on this concept of translatability and on how it correlates with post-editing effort. It draws on research that is ongoing at the time of writing.[1] In section 2 the notions of translatability and translatability measurement are presented. In section 3, methods for measuring post-editing effort are explained. Section 4 describes the research methodology in more detail. Section 5 presents data derived from the work-in-progress where the correlations between post-editing and translatability are being explored. A summary and future work will be outlined in Section 6.

2. Source Text Translatability

Several researchers list "translatability" as one of the main goals of controlled language (Wojcik and Hoard 1996; Reuther 1998; Means and Godden 1996). To date, five researchers have written specifically on translatability measurement in the domain of CL, i.e. (Gdaniec 1994; Bernth 1999a, 1999b; Bernth & McCord 2000; Underwood and Jongejan 2001; Bernth and Gdaniec

[1] The research was funded by the Irish Research Council for the Humanities and Social Sciences as well as by the School of Applied Language and Intercultural Studies at Dublin City University.

2002. The approaches to translatability measurement outlined by these researchers are broadly similar: linguistic features known to be problematic for MT (called "negative sentence properties" or "Translatability Indicators", henceforth "TIs") are identified and pre-determined numeric penalties are applied to them. Examples of such linguistic features include, but are not limited to, the passive voice, the gerund, long noun phrases and ellipsis.[2] In previous studies on translatability, it has been acknowledged that some TIs are more problematic for certain language pairs and directions than others. In addition, some machine translation systems can cope better with specific TIs than others. For both of these reasons, weights are applied to the penalty values for TIs. A value for the overall machine translatability of the source text is calculated by combining the penalty values and weights.

The numeric values used in the research mentioned above differ from one researcher to the next. However, it is not the values that matter, but their relative significance. For example, if a source text in English, which is to be translated into German by MT system X scores 5 out of 10, what does that actually tell us about the suitability of that source text for machine translation into that target language? To date, researchers have used quality evaluation of the target-text to correlate the values of penalties with MT quality. However, it has been acknowledged that MT quality evaluation is a highly complex, and often subjective, task (King et al,

[2] A detailed discussion of TIs is beyond the scope of this paper. In general, it can be said that TIs are what Machine-Oriented Controlled Language (MOCL) rules seek to identify and eliminate. For a more detailed discussion on MOCL rules in English, see (O'Brien: 2003). For more information on TIs, see (Gdaniec 1994; Bernth 1999a, 1999b, Bernth & McCord 2000; Underwood and Jongejan 2001; Bernth and Gdaniec 2002).

2003). In addition, the relationship between quality evaluation and post-editing effort has not yet, to my knowledge, been explored in detail. The core motivation for the research described in this paper lies in the pursuit of an answer to the question: what are the correlations, if any, between source text translatability and post-editing effort?

3. How Can Post-Editing Effort Be Measured?

The task of a post-editor, is to "edit, modify and/or correct pre-translated text that has been processed by a machine translation system from a source language into (a) target language(s)" (Allen: 2003). According to several authors, the activity of post-editing differs from traditional translation and from revision, i.e. reviewing text translated by a human and correcting errors such as content, spelling, punctuation, formatting etc. (McElhaney and Vasconcellos, 1988; Piggott 1988; Loffler-Laurian 1986). Allen (2003) states that "very little concrete data has been made available with regard to user studies, results and the methodology employed for post-editing". The process of post-editing has been described in some detail by Rossi (1982), Wagner (1985), Loffler-Laurian (1986), Senez (1998a), Allen (2003), amongst others. However, one description of the post-editing process that far outweighs others in its level of detail is that of Krings (2001).

According to Krings "the question of post-editing effort is the key issue in the evaluation of the practicality of machine translation systems" (2001: 178). He maintains that post-editing effort can be measured on three levels: temporal, technical and cognitive. Temporal effort obviously refers to the time taken to post-edit a sentence to a particular level of quality. Technical effort refers to deletions, insertions and text re-ordering. Cognitive effort refers to the extent and type of cognitive processes that must be activated to remedy a deficiency in MT output. Temporal and technical effort can be directly observed and it is reasonable to assume that they are linked to cognitive effort. Unlike temporal and technical effort, cognitive effort cannot be observed directly. Consequently, in order to observe cognitive effort Krings used the method known as "Think-Aloud Protocol" (TAP).

Think-Aloud Protocol requires a research subject to verbalise what is going through his or her mind while s/he is completing a task. It was first used as a research method in the domain of psychology. Think-Aloud Protocol has enjoyed a certain level of popularity in translation process research.[3] However, Krings's observations on the influence of TAP on the post-editing process strongly suggest that it is inappropriate as a method for the research project described here. First and foremost, Krings discovered that the use of TAP slowed down the post-editing process by almost 33%, when compared with a group who did not use TAP. Secondly, he discovered that TAP increased the number of non-linear writing acts (e.g. deletions, insertions etc.). Since the focus of the current study is the correlation, if any, between temporal, technical and cognitive post-editing effort and source-text translatability, it is clear from these two conclusions alone that TAP is not a suitable methodology.[4]

---

[3] For example see (Hansen 2003; Alves 2003; Jääskelainen 1987, 1989; Séguinot 1989; Lörscher 1991).

[4] TAP is fraught with other difficulties, e.g. interference with the task at hand, incompleteness of data, and varying levels of individual verbalisation willingness. For more on this topic, see

Keyboard monitoring is one alternative methodology for recording and measuring post-editing effort. *Translog* is a keyboard monitoring tool which logs electronic text production. It was developed at the Copenhagen Business School specifically for the monitoring of translation activity within the framework of translation process research (Hansen, 1999).[5]

Translog is made up of two applications: Translog User and Translog Supervisor. Translog Supervisor is used to set-up projects while Translog User is used to translate texts and to record text processing activity during translation. Once a project has been set up in Translog Supervisor, it is then opened in Translog User and the source text is translated. The Translog User interface is presented in Figure 1:

(Tirkkonen-Condit 1991; Lörscher 1989; Séguinot 1996).

[5] Thanks to Arnt Lykke Jakobsen, developer of Translog, who gave permission to use Translog in this research and who helped with many queries.

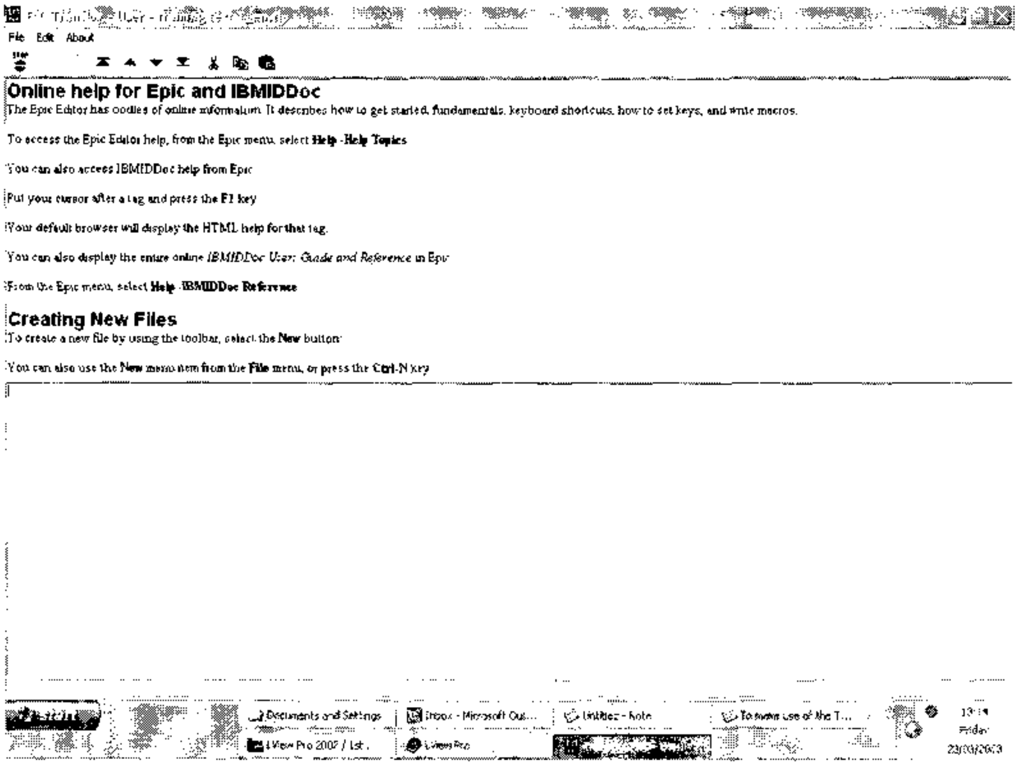

Figure 1: Translog User Interface

When the translator is ready to start a task, the green flag on the toolbar is clicked and this prompts Translog User to commence the logging process. While working, the translator can write, delete, move, copy, cut and paste text. S/he can also look a term up in the Translog dictionary or scroll through the source and target texts. When the translation is complete, the "Stop" sign is clicked and the log file is saved.

In order to monitor post-editing activity, the researcher must first paste the MT output into the target text window before the translator can start post-editing (see Figure 2).

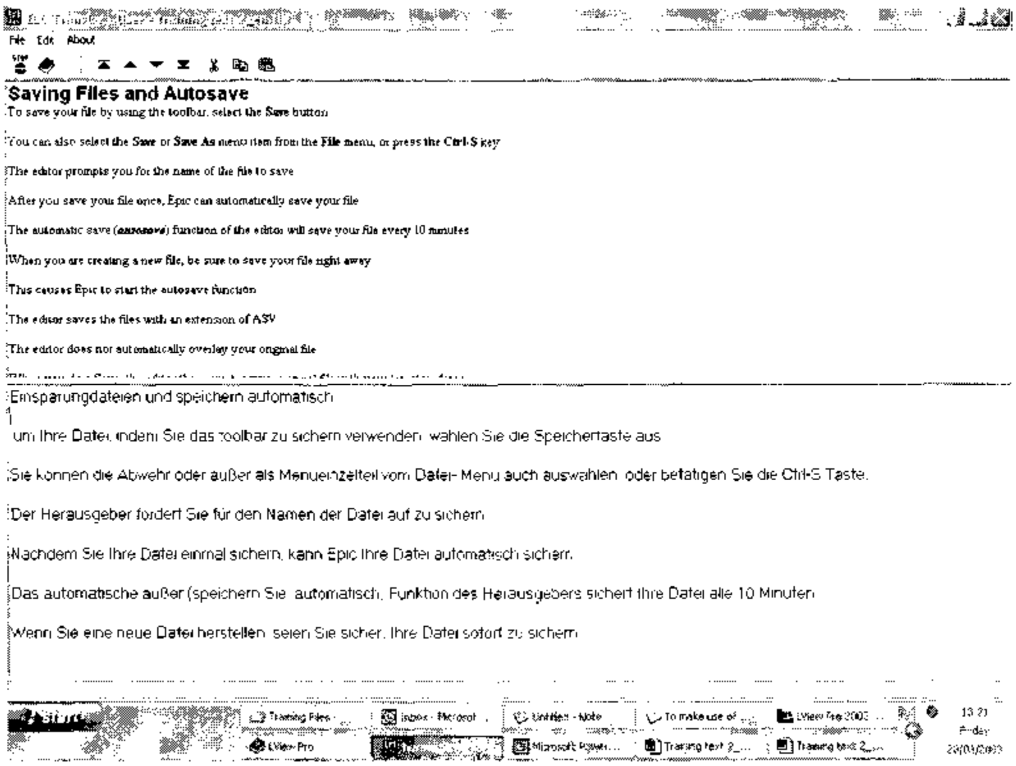

Figure 2: Translog User - Post-Editing Scenario

Translog produces a log file, which is a linear representation of all keyboarding activity, e.g. letters typed or deleted, pauses and their duration, cut and paste activity, mouse activity and electronic dictionary look-up.[6] The log file can then be used to replay the translation or post-editing process at user-defined speeds. In effect, Translog produces both a "photograph" and a "video" of the translation or post-editing process.

For this research project, Translog has been used to collate data on the temporal, technical and cognitive effort involved in post-editing. Temporal data are collated by recording the time required to post-edit individual sentences.

[6] "Dictionary look-up" refers only to the use of Translog's in-built dictionary and not to the use of other electronic or paper dictionaries.

Technical data are collated by recording the number of words inserted, words deleted, the number of cut and paste actions and dictionary look-ups. Cognitive effort, as Krings states, cannot be observed directly. However, it has been suggested that pauses are good indicators of cognitive effort and Translog allows us to observe the number, location and duration of pauses.[7]

In order to triangulate results, a second methodology was chosen for measuring post-editing effort and, in particular, cognitive effort. That methodology is called "Choice Network Analysis" (CNA) (Campbell 1999, 2000a, 2000b) (Campbell and Hale 1999). In Campbell (2000b), CNA is explained in the following manner:

> .. .Choice Network Analysis compares the renditions of a single string of translation by multiple translators in order to propose a network of choices that theoretically represents the cognitive model available to any translator for translating that string. The technique is favoured over the think-aloud method, which is acknowledged as not being able to access automaticized processes.
>
> (ibid: 215)

Thus CNA is presented as a method for constructing models of the mental processing underlying translation and it is also useful for estimating the relative difficulty of parts of source texts, where the measure of difficulty can be established based on the complexity of choices available to the translator.

CNA is best explained by example. The following sentence is taken from

[7] For more on the topic of pauses as cognitive indicators, see (Krings, 2001: 304; Séguinot, 1989b: 31; Schilperoord, 1996). For more on the topic of Translog and how it records pauses, see (Jakobsen 1998).

the source text used in the current research:

> The IBMIDDoc document type definition defines the default structure for all IBM documentation.

The machine translation output for this sentence in German is:

> Die IBMIDDoc Dokumenttypendefinition definiert die Standardstruktur für alle IBM Dokumentation.[8]

Table 1 provides us with a snapshot of the MT output compared with the post-edited versions of this sentence by nine post-editors (referred to as P1-P9). The text marked with bold highlights the areas where most post-editing activity occurred.[9]

[8] The author would like to thank Arendse Bernth and Ulrich Heess of IBM for their support in providing materials, technology and professional translators for this research.

[9] Square brackets indicate where the text contained in square brackets occurred in the post-edited version.

Table 1: MT Output vs. nine post-edited versions[10]

MT	Die IBMIDDoc	**Dokumenttypendefinition**	definiert	die Standardstruktur	für	**alle**	IBM Dokumentation.	-
P1	Die IBMIDDoc-	**Dokumenttypendefinition**	definiert	die Standardstruktur	fur	die **gesamte**	IBM Dokumentation.	-
P2	Die IBMIDDoc-	**Dokumenttypdefinition**	definiert	die Standardstruktur	für	**alle**	IBM Dokumentationen.	-
P3	Die IBMIDDoc	**Dokumenttypendefinition**	definiert	die Standardstruktur	für	**jegliche**	IBM Dokumentation.	-
P4	Die IBMIDDoc	**Dokumenttypendefinition**	definiert	die Standardstruktur	für	die **gesamte**	IBM Dokumentation.	-
P5	[IBMIDDoc]	**Die Definition des Dokumenttyps [...]**	definiert	die Standardstruktur	für	alle	IBM Dokumentationen,	-
P6	Die IBMIDDoc	**Dokumenttypendefinition**	definiert	die Standardstruktur	fü	die **gesamte**	IBM Dokumentation.	-
P7	[IBMIDDoc]	**Die Dokumenttypendefinition** [...]	definiert	die Standardstruktur	für	**alle**	IBM Dokumentationen.	-
P8	Die IBMIDDoc	**Dokumenttypendefinition**	definiert	die Standardstruktur	für	die **gesamte**	IBM Dokumentation.	-
P9	Die IBMIDDoc-	**Dokumenttypdefinition**	legt	die Standardstruktur	- -	der **gesamten**	IBM Dokumentation	fest

Table 1 shows that all post-editors focused on two features in the MT output, i.e. *Dokumenttypendefinition* and *alle.* By applying the theory of CNA, we can conclude that the long noun phrase (NP) *document type definition* and the quantifier *all* are the two features in the source text that represent difficulty. In other words, they are the two TIs.

In this section we posed the question "How can post-editing effort be measured?" The use of TAP as a method for measuring post-editing effort was discounted due to the numerous drawbacks mentioned by Krings. Keyboard

[10] It should be noted that this table format is not how Campbell and Hale represent networks in their publications. The author found the representation of text chunks in tables easier to analyse.

monitoring using Translog and Choice Network Analysis were presented as alternative methods. In section 5 we will implement these methods to explore the potential correlations between post-editing effort and source text translatability. Prior to that, we will describe in more detail how the methodological tools described above are put to use.

4. Methodology

The source document used in this research is an excerpt from a manual describing an SGML-based editor. The number of words totals 1 775. The TIs in the source text were identified using a controlled language application.[11] The source text was then edited so that the TIs identified in Bernth and McCord 2000, who provide the most detailed list of TIs in the literature on CL, occurred at least twice in the document.[12] The edit phase also aimed at producing some sentences that had "minimal TIs", i.e. none of the indicators listed in the literature occurred in some sentences. This was to enable a comparison of post-editing effort across sentences with TIs and sentences with minimal TIs. The document was then analysed by the MT system for words that did not exist in the dictionary and the dictionary was coded with these words. Finally, the document was machine translated using IBM WebSphere English-German. Twelve professional translators, who were native speakers of German and who had extensive translation experience in both this text type and domain, were recruited for

[11] The IBM CL tool "EasyEnglishAnalyzer" was used for this purpose.

[12] Given the restrictions on time and resources, two occurrences of each translatability indicator was the maximum number we could allow.

post-editing. The translators' profiles were as homogenous *as* possible: all translators were regular users of a translation memory and term management application but they had little or no experience of machine translation and post-editing; Insofar as possible, they had similar professional qualifications and they had similar numbers of years of translation experience. The translators participated in a short training session on Translog before commencing the post-editing task. They were instructed to post-edit the MT output to a level that was accurate and acceptable to a native speaker of German. They were also instructed to complete only one pass through the text. Although this deviates from a "normal" working scenario, it was necessary to set the experiment up in this way so that the correlation between translatability and post-editing effort could be examined.[13] Subjects were also advised not to make any unessential, stylistic changes. For comparative purposes, a random selection of three out of the twelve post-editors were asked to translate, rather than post-edit, the document. All subjects were given two hours to complete the task and all completed within that time-frame.

The first stage of data analysis involves the systematic application of the Choice Network Analysis model to each sentence.[14] This provides us with an indication as to which parts of the source sentence gave rise to post-editing effort. That information is then analysed against the known TIs in the sentence in order to ascertain which TIs actually gave rise to post-editing effort.

---

[13] It was surmised that if translators knew that they could revise the text, they may not have put in the same "effort" during the first pass.

[14] The term "sentence" is used here to refer to a chunk of text. It includes not only standard sentences, but also bulleted list items, headings, figure labels etc.

Following this, the post-editing of each sentence by the nine post-editors is replayed in Translog. The following data categories are then recorded:

- Number of source words
- Total time required for post-editing (in seconds)
- Processing speed (i.e. number of words processed per second)
- Time spent pausing
- Pause ratio (i.e. percentage of total processing time spent pausing)
- Number of words deleted
- Number of words inserted
- Number of cut actions
- Number of paste actions
- Number of successful dictionary look-ups (i.e. word found in Translog's dictionary)
- Number of unsuccessful dictionary look-ups (i.e. word not found in Translog's dictionary)
- Relative post-editing effort

The median value for each of these categories is also recorded. The "relative post-editing effort" is automatically calculated by applying a formula suggested by Krings (2001). This measure gives an indication of the post-editing effort vis-a-vis the translation effort for the same sentence. It is calculated by dividing the average processing time for post-editing by the average processing time for translation. A value between 0 and 1 means that post-editing effort was less than translation effort. A value of 1 means that translation and post-editing effort are equal and a value greater than 1 means that post-editing is more time-consuming than translation.[15] Figure 3 shows the relative post-editing effort for the first 40 sentences.

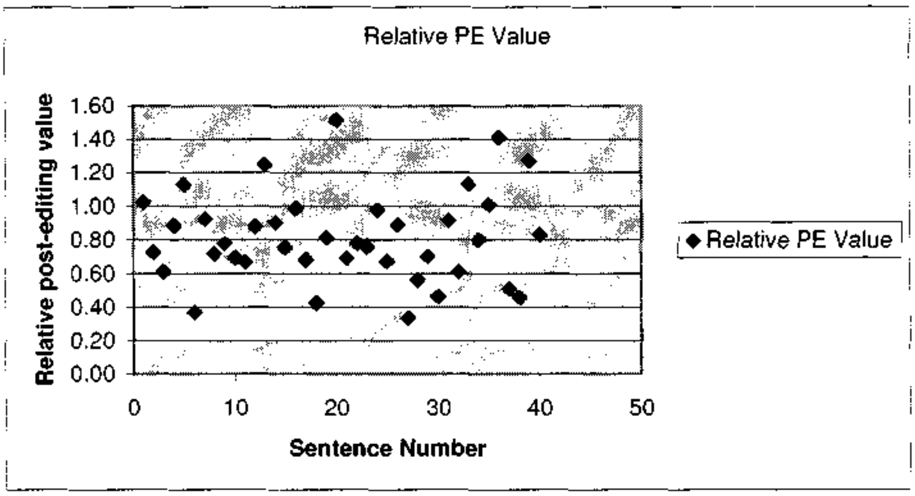

Figure 3: Relative Post-editing effort

We can see from the data in Figure 3 that eight out of the 40 sentences (i.e. 20%) lie above the value of 1, indicating that post-editing effort was greater than translation effort for these sentences. The remainder lie below 1, indicating that post-editing effort was lower than translation effort for these sentences. Nonetheless, it is interesting to note that most of the sentences cluster between the values of 0.60 and 1. This suggests that post-editing effort is close to translation effort for the majority of sentences.

A final stage in the analysis involves an investigation of pause activity. Pauses recorded in the Translog linear representation file are isolated. Particular

[15] For a more detailed discussion of this measurement, see (Krings, 2001:182)

attention is paid to what occurs after each pause, e.g. a deletion, an addition, a cursor movement, and attempts are made to triangulate these data with the Choice Network Analysis data.[16]

5. Correlations Between Post-Editing Effort and Source Text Translatability

The methodology described in the previous section has so far been applied to 40 out of 165 sentences. Seven out of these 40 sentences can be described as containing "minimal TIs", i.e. none of the TIs listed in Bernth and McCord (2000) are contained in these sentences. The median processing speeds over nine post-editors for these sentences have been calculated. Processing speed is defined as the number of source words processed per second and it is calculated by dividing the number of source words in a sentence by the total processing time in seconds. Figure 4 compares the median processing speed for sentences with TIs against the median processing speed for the seven sentences with minimal TIs.

[16] Although it has been suggested that pauses are indicators of cognitive processing (Schilperoord 1996), the current research suggests that the analysis of pauses and subsequent editing activity fails to provide a *comprehensive* account of cognitive processing since there is a high occurrence of pauses that are followed only by cursor and/or mouse movements and not by any editing activity. It is reasonable to assume that cognitive effort is being expended during these pauses, but the content of those pauses are not available to us for analysis. An in-depth discussion of this problem is beyond the scope of the current paper, but it is the author's intention to address this problem in more detail in the future.

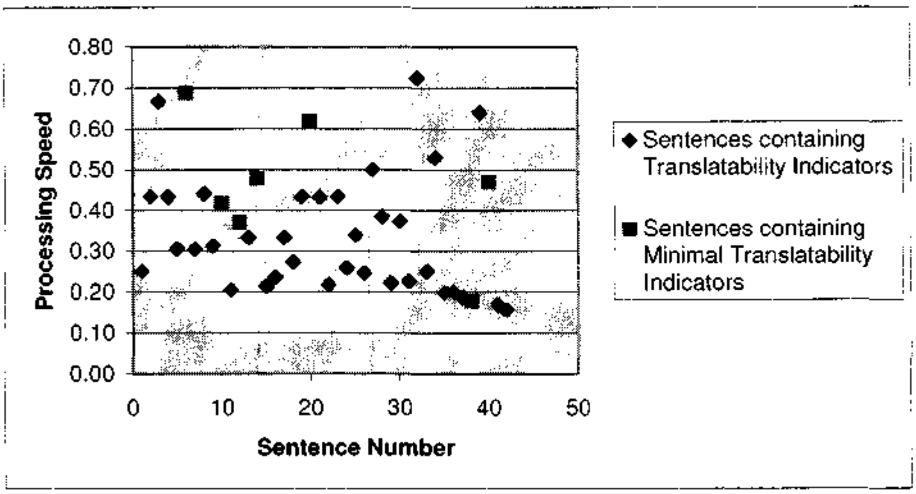

Figure 4: Median processing speeds: sentences with TIs vs. sentences with minimal TIs

While there are some exceptions, we can observe a trend whereby the sentences with minimal TIs are processed more quickly on average, when compared with those sentences with TIs.

As previously mentioned, one of the ways of calculating technical post-editing effort is by observing the number of words, or parts of words, deleted and inserted. Figure 5 shows us the median number of words (or parts of words) deleted for sentences containing TIs as against those sentences with minimal TIs. Figure 6 shows us the number of words (or parts of words) inserted for the same.

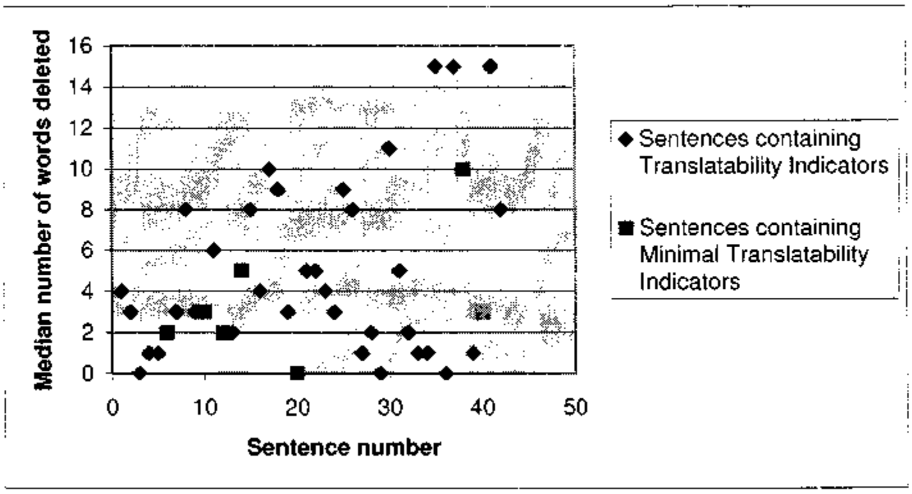

Figure 5: Median number of words deleted: TIs vs. minimal TIs

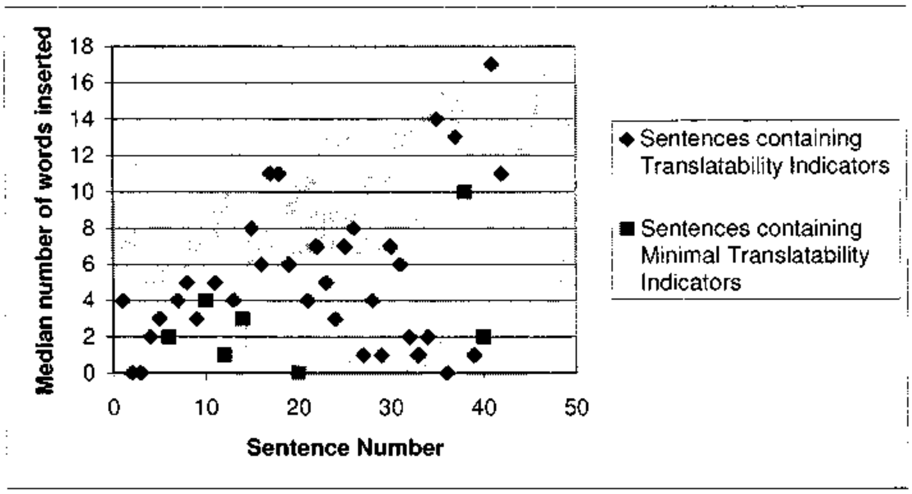

Figure 6: Median number of words inserted: TIs vs. minimal TIs

The first observation we can make based on these figures is that deletions and insertions appear to follow a similar trend. While there are some exceptions, deletions and insertions, i.e. technical post-editing effort, occur, on average, more frequently in sentences with TIs compared to sentences with minimal TIs. It is interesting to note, however, that technical post-editing effort in the form of deletions and insertions is still required to some extent for sentences with minimal TIs and that the level of such effort is sometimes equal to sentences containing TIs.

We can draw some tentative conclusions from the figures above regarding post-editing effort and translatability indicators: processing speed appears to be faster for sentences with minimal TIs and technical post-editing effort in the form of deletions and insertions appears to be lower for such sentences. Also, in some cases, sentences with minimal TIs still require post-editing effort.

In order to learn more about the correlations between specific TIs and post-editing effort, three questions have been posed:

(1) Of the sentences analysed to date, which TIs appear to consistently create post-editing effort?

(2) Of the sentences analysed to date, which TIs appear NOT to create post-editing effort?

(3) Of the sentences with minimal TIs analysed to date, which linguistic features appear to create unexpected post-editing effort?

One TI that consistently causes post-editing effort in our data is the long noun phrase (NP). Long NPs occur four times in our sentences analysed to date. Figure 7 compares the processing speed for sentences containing long NPs with sentences with minimal TIs. These figures clearly show a higher temporal post-editing effort for sentences containing long noun phrases:

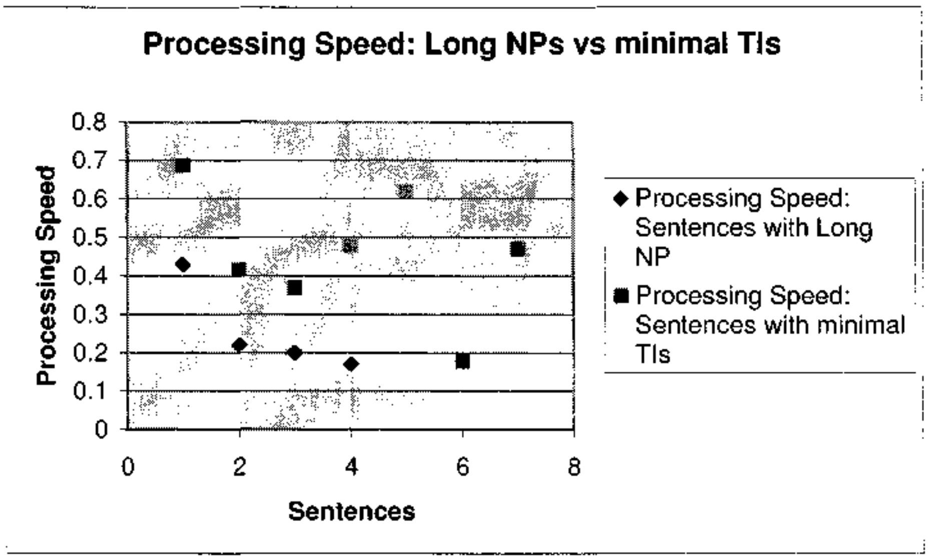

Figure 7: Processing speed for sentences with long NPs vs. sentences with minimal TIs

Figures 8 and 9 compare the median numbers of words inserted and deleted for sentences containing long NPs and those with minimal TIs.

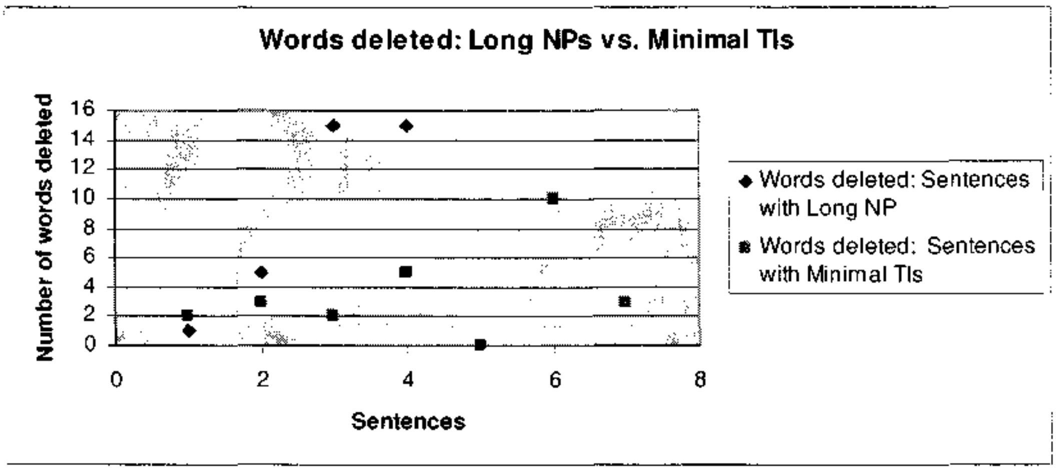

Figure 8: Words deleted in sentences with long NPs vs. sentences with minimal TIs

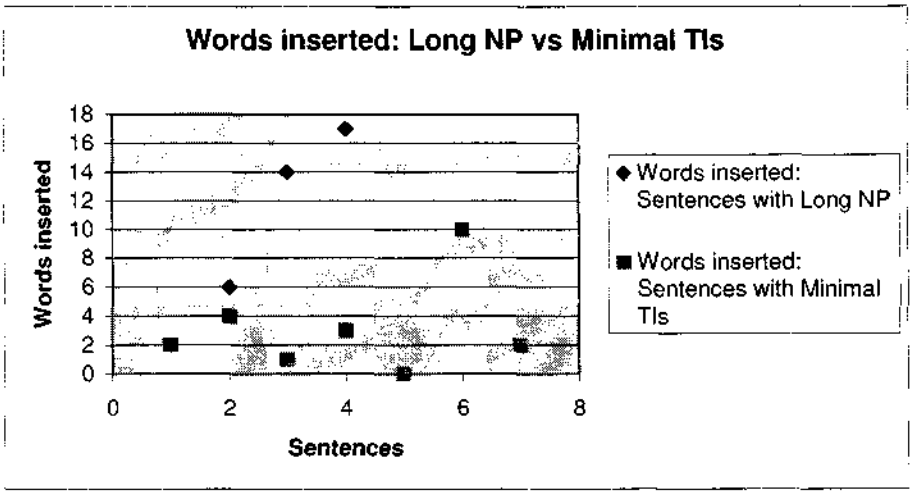

Figure 9: Words inserted in sentences with long NPs vs. sentences with minimal TIs

In both cases, we have two sentences where the number of words inserted or deleted are similar to the figures for sentences with minimal TIs, i.e. they cluster between 0 and 4. However, we also have two sentences where the numbers of words inserted or deleted are significantly higher for sentences with TIs, i.e. between 14 and 18.

A similar trend can be seen when we examine the same indicators of post-editing effort for the gerund in English. Figures 10, 11 and 12 show median processing speeds, numbers of words inserted and deleted for sentences containing the gerund compared with sentences with minimal TIs.

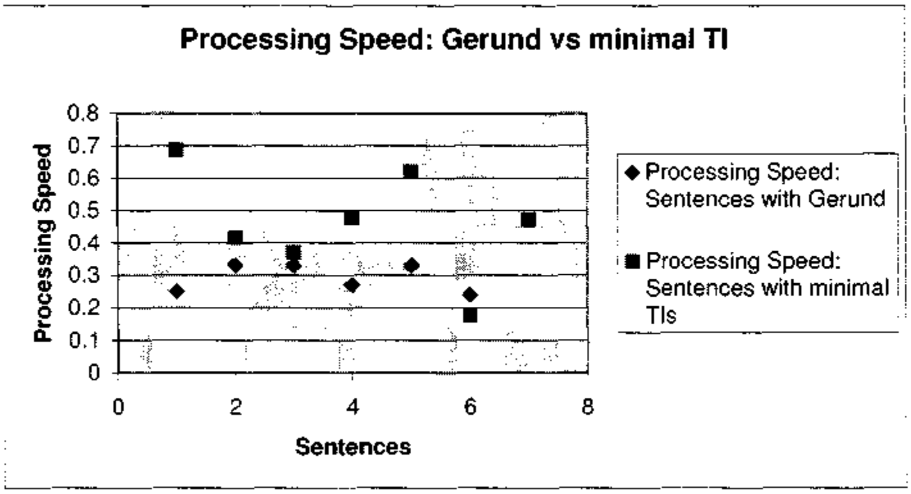

Figure 10: Processing speed: Gerund vs. minimal TIs

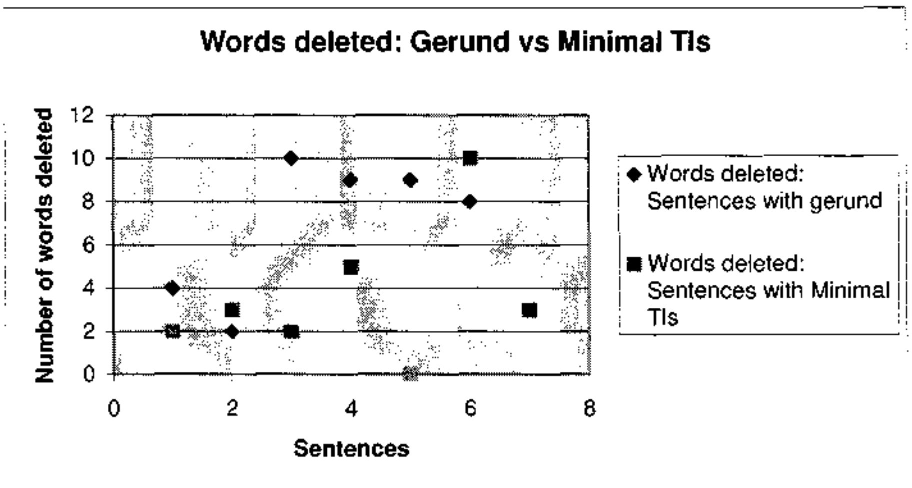

Figure 11: Words deleted: Gerund vs. minimal TIs

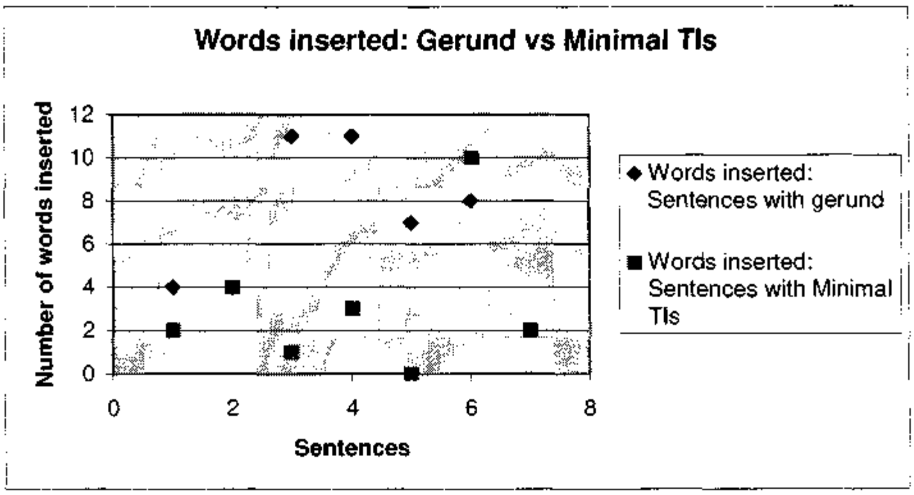

Figure 12: Words inserted: Gerund vs. Minimal TIs

On average, we can say that the processing speed is slower for sentences containing the gerund in English and the number of words deleted and inserted is higher.

We can see from the data above that the post-editing effort for sentences with minimal TIs is not *consistently* lower than that of sentences with TIs. The processing speeds recorded for sentence 36 are lower than for other minimal TI sentences. Also, the number of words inserted and deleted is more in line with the values for sentences containing TIs. Choice Network Analysis combined with the playback feature in Translog reveal the reasons for this. The source sentence reads:

See Generated Text for more on this feature.

The MT output was:

Sehen Sie generierten Text für mehr auf diesem Produktmerkmal.

CNA reveals that all post-editors were unhappy with this MT output and all decided to change it. An example of one post-edited version is:

Weitere Informationen zu dieser Funktion finden Sie im Abschnitt "Generierter Text".

The final post-edited product is significantly different from the MT output, indicating that post-editing effort was high. This is one of three sentences where the structure "*See X for more information"* occurs*.* Each occurrence of this phrase was post-edited to a significant extent by all post-editors.

The data analysed to date suggest that, in addition to long noun phrases and the gerund, the following TIs produce MT output that requires post-editing: slang, ellipsis, non-finite verbs, misspelling and problematic punctuation. Interestingly, the data also suggest that some TIs do not result in post-editing effort, e.g. proper nouns, abbreviations, and a missing *in order to* phrase.[17] Time and space restrictions mean that we cannot present graphs for all of these features here.

6. Conclusions and Future Work

The results presented here have been produced using a sub-set of the total number of sentences and TIs under analysis. Nonetheless, some tentative

[17] We would suggest that when proper nouns and abbreviations are contained in an MT dictionary, they should not figure as TIs.

findings can be highlighted and it is expected that analysis of the remaining sentences and TIs will further underpin the findings presented here.

The data suggest that the post-editing speed for sentences with minimal TIs is, on average, faster than for sentences with TIs. This implies that the application of CL rules to improve machine translatability will indeed reduce post-editing effort.

However, the data also suggest that some TIs, e.g. long NPs and the gerund, cause post-editing effort while others, e.g. abbreviations and proper nouns, do not. Further analysis will allow us to specify exactly which TIs are most problematic. It is hoped that this kind of information could be used in the tuning of CL rules in the future.

Finally, although a sentence may not contain a TI, the data suggest that such a sentence may still be subject to post-editing activity. One of the aims of the ongoing research is to identify the reason(s) why such sentences are being post-edited.

References

Allen, Jeff (2003), "Post-editing". Harold Somers, ed. *Computers and Translation: A translator's guide.* Amsterdam-Philadelphia: John Benjamins. 297-318.

Alves, Fabio ed. (2003), *Triangulating Translation.* Amsterdam-Philadelphia: John Benjamins.

Bernth, Arendse (1999a), "A Confidence Index for MT", in *Proceedings of the 8th International Conference on Theoretical and Methodological Issues in Machine Translation,* Chester, 120-127.

Bernth, Arendse (1999b), "Controlling Input and Output of MT For Greater User Acceptance", in Proceedings of the 21st Conference on Translating and the Computer, ASLIB, London.

Bernth, Arendse and Claudia Gdaniec (2002), "MTranslatability". *Machine Translation* Vol 16, No. 3. 175-218.

Bernth, Arendse and Michael McCord (2000), "The effect of source analysis on translation confidence". *Proceedings of AMTA-2000.* Mision del Sol, Cuernavaca, Mexico, October 10-14. 89-99.

Campbell, Stuart (1999), "A cognitive approach to source text difficulty in translation". *Target* 11:1. 33-63.

Campbell, Stuart (2000a), "Choice Network Analysis in Translation Research". Maeve Olohan, ed., *Intercultural Faultlines: Research models in translation studies: textual and cognitive aspects,* St. Jerome, 29-42.

Campbell, Stuart (2000b), "Critical structures in the evaluation of translations from Arabic into English as a second language". *The Translator* Vol. 6, number 2. 211-229.

Campbell, Stuart and Sandra Hale (1999), "What makes a text difficult to translate?". *Proceedings of the 1998 ALAA Congress.* Available at: http://www.cltr.uq.edu.au/alaa/proceed/camphale.html. Last checked on:

April 6, 2004.

EAMT/CLAW (2003), Joint conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop, 15th-171h May, Dublin City University.

Gdaniec, Claudia (1994), "The Logos translatability index". *Proceedings of the First Conference of the Association for Machine Translation in the Americas.* Columbia, Maryland, 5-8 October. 97-105.

Hansen, Gyde (1999), *Probing the process in translation: methods and results.* Copenhagen Studies in Language. Copenhagen: Samfundslitteratur.

Hansen, Gyde (2003), "Controlling the process: theoretical and methodological reflections on research into translation processes". Fabio Alves, ed. *Triangulating Translation.* Amsterdam-Philadelphia: John Benjamins. 25-42.

Jääskelainen, Riitta (1987), "What happens in a translation process? - Think-aloud protocols of translation", Unpublished *pro gradu* study, Savonlinna School of Translation Studies.

Jääskelainen, Riitta (1989), "Translation assignment in professional vs. non-professional translation: A think-aloud protocol study", in Candace Seguinot (ed) *The Translation Process,* Toronoto: H.G. Publications, School of Translation, York University, 87-98.

Jakobsen, Arnt Lykke (1998), "Logging time delay in translation", in *LSP texts and the translation process,* Copenhagen Working Papers, 73-101.

King, Margaret, Andrei Popescu-Belis and Eduard Hovy (2003), "FEMTI: Creating and using a framework for MT evaluation", in Online proceedings of MT summit IX, New Orleans, http://www.amtaweb.org/summit/MTSummit/papers.html (last checked: October 10, 2004.

Krings, Hans Peter (2001), *Repairing Texts: empirical investigations of machine translation post-editing processes.* G.S. Koby, ed. Kent, Ohio: Kent State University Press.

Loffler-Laurian, Anne-Marie (1986a), "Post-édition rapide et post-édition conventionelle: deux modalités d'une activité specifique I". *Multilingua* 5:4.81-88.

Loffler-Laurian, Anne-Marie. (1986b), "Post-édition rapide et post-édition conventionelle: deux modalités d'une activite specifique II". *Multilingua* 5:4. 225-229.

Lörscher, Wolfgang (1991), *Translation performance, translation process, and translation strategies: a psycholinguistic investigation.* Tübingen: Gunter Narr Verlag.

McElhaney, Terrence and Muriel Vasconcellos. 1988. "The translator and the post-editing experience." Muriel Vasconcellos, ed. *Technology as translation strategy.* American Translators Association Scholarly Monograph Series, Vol II. Binghamton: State University of New York, SUNY. 140-148.

Means, Linda & Kurt Godden (1996), "The Controlled Automotive Service Language (CASL) Project", in Geert Adriaens et al (eds), *Proceedings of the First International Workshop on Controlled Language Applications (CLAW 96),* Centre for Computational Linguistics, Leuven, Belgium.

Mitamura, Teruko, Eric Nyberg, Geert Adriaens, Linda Schmandt, Rick Wojcik, Remi Zajac (eds) (1998), *Proceedings of the Second International Workshop on Controlled Language Applications, Language Technologies Institute,* Carnegie Mellon University, Pittsburgh, Pennsylvania.

Nyberg, Eric, Teruko Mitamura and Willem-Olaf Huijsen (2003), "Controlled language for authoring and translation". Harold Somers, ed. *Computers and Translation: A translator's guide.* Amsterdam-Philadelphia: John

Benjamins. 245-282.

O'Brien, Sharon (2002), "Teaching post-editing: A proposal for course content". *Teaching Machine Translation - the 6th International Workshop of the European Association of Machine Translation.* November 14-15, Centre for Computational Linguistics, UMIST, Manchester, England. 99-106.

O'Brien, Sharon (2003), "Controlling controlled English: An analysis of several controlled language rule sets", in EAMT/CLAW (2003), Joint conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop, 15th-17th May, Dublin City University, 105-114.

Piggott Ian, (1988), "MT in large organizations: Systran and the Commission of the European Communities". Muriel Vasconcellos, ed. *Technology as translation strategy.* American Translators Association Scholarly Monograph Series, Vol II. Binghamton: State University of New York, SUNY. 159-166.

Reuther, Ursula (1998), "Controlling Language in an Industrial Application", in Mitamura et al (eds), 174-184.

Rossi Francesco, (1982), "The impact of post-editors' feedback on the quality of MT", Veronica Lawson, ed. *Practical Experience of Machine Translation: Proceedings of a conference.* London, 5-6 November. 113-118.

Schilperoord, J (1996), *It's about time: Temporal aspects of cognitive processes in text production,* Amsterdam: Rodopi.

Séguinot, Candace (1989), "The translation process: An experimental study", in Candace Séguinot (ed) *The Translation Process,* Toronto: H.G. Publications, School of Translation, York University, 21-53.

Séguinot, Candace (1996), "Some thoughts about think-aloud protocols". *Target* 8:1.75-95.

Senez, Dorothy (1998), "Post-editing service for machine translation users at the European Commission". *Proceedings of Translating and the Computer 20.* London: ASLIB.

Tirkkonen-Condit, Sonja (1991), "Empirical research in translation and intercultural studies; selected papers of the TRANSIF seminar", Savonlinna.

Underwood, Nancy & B. Jongejan (2001), "Translatability checker: A tool to help decide whether to use MT". Bente Maegaard, ed. *Proceedings of the MT Summit VIII: Machine translation in the information age.* 18-22 September, Santiago de Compostela, Spain. 363-368.

Van der Meer, Jaap (2003), "At last translation automation becomes a reality: An anthology of the translation market" in EAMT/CLAW (2003), Joint conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop, 15th-17th May, Dublin City University, 180-184.

Wagner, Elizabeth.(1985), "Rapid Post-editing of Systran". Veronica Lawson, ed. *Tools for the trade: Translating and the computer 5.* London: ASLIB. 199-213.

Wojcik, Rick & James Hoard (1996), "Controlled languages in industry", in Ronald Cole et al (eds) *Survey of the state of the art in human language technology,* http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html (last checked: October 10th, 2004).